LongRunMIP - motivation and design for a large collection of millennial-length AO-GCM simulations

Article

Accepted Version

LongRunMIP – motivation and design for a large collection of millennial-length AO-GCM simulations

Maria Rugenstein*

Institute for Atmospheric and Climate Science, ETH Zurich, CH-8092 Zurich, Switzerland

Max-Planck-Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany

Jonah Bloch-Johnson

NCAS, University of Reading, Reading

Ayako Abe-Ouchi

Atmosphere and Ocean Research Institute, The University of Tokyo

Timothy Andrews

Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB

Urs Beyerle

Institute for Atmospheric and Climate Science, ETH Zurich, CH-8092 Zurich, Switzerland

Long Cao

School of Earth Sciences, Zhejiang University, Hang Zhou, Zhejiang Province, 310027, China

Tarun Chadha

ITS Research Informatics, ETH Zurich, Switzerland

Gokhan Danabasoglu

National Center for Atmospheric Research, P.O. Box 3000, Boulder, CO 80307

Jean-Louis Dufresne

Centre National de la Recherche Scientifique, Universit Pierre et Marie Curie, ENS, Ecole Polytechnique, Paris, France

Lei Duan

School of Earth Sciences, Zhejiang University, Hang Zhou, Zhejiang Province, 310027, China

Marie-Alice Foujols

Institut Pierre-Simon-Laplace, Sorbonne Universit / CNRS, Paris, France

Thomas Frölicher

Climate and Environmental Physics, Physics Institute, University of Bern, Switzerland, Oeschger Centre for Climate Change Research, University of Bern, Switzerland

Olivier Geoffroy

CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Jonathan Gregory

NCAS, University of Reading, Reading, Met Office Hadley Centre, FitzRoy Road, Exeter, EX1 3PB

Reto Knutti

Institute for Atmospheric and Climate Science, ETH Zurich, CH-8092 Zurich, Switzerland

Chao Li

Max-Planck-Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany

Alice Marzocchi

National Oceanography Centre, European Way, Southampton, SO14 3ZH, UK

Thorsten Mauritsen

Stockholm University, SE-106 91 Stockholm, Sweden

Matthew Menary

LOCEAN, Sorbonne Universit, Paris, France

Elisabeth Moyer

Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois, USA

Larissa Nazarenko

NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025

David Paynter

Geophysical Fluid Dynamics Laboratory, Princeton, New Jersey, USA

David Saint-Martin

CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

Gavin A. Schmidt

NASA Goddard Institute for Space Studies, 2880 Broadway, New York, NY 10025

Akitomo Yamamoto

Japan Agency for Marine-Earth Science and Technology, Yokohama, Japan

Shuting Yang

Danish Meteorological Institute, Lyngbyvej 100, DK-2100 Copenhagen, Denmark

**Corresponding author address:* Maria Rugenstein, Max-Planck-Institute for Meteorology, Bundesstrasse 53, 20146 Hamburg, Germany

E-mail: maria.rugenstein@mpimet.mpg.de

ABSTRACT

We present a model intercomparison project, LongRunMIP, the first collection of millennial-length (1000+ year) simulations of complex coupled climate models with a representation of ocean, atmosphere, sea ice, and land surface, and their interactions. Standard model simulations are generally only a few hundred years long. However, modeling the long-term equilibration in response to radiative forcing perturbation is important for understanding many climate phenomena, such as the evolution of ocean circulation, time- and temperature-dependent feedbacks, and the differentiation of forced signal and internal variability. The aim of LongRunMIP is to facilitate research into these questions by serving as an archive for simulations that capture as much of this equilibration as possible. The only requirement to participate in LongRunMIP is to contribute a simulation with elevated, constant CO_2 forcing that lasts at least 1000 years. LongRunMIP is a MIP of opportunity in that the simulations were mostly performed prior to the conception of the archive without an agreed-upon set of experiments. For most models, the archive contains a preindustrial control simulation and simulations with an idealized (typically abrupt) CO_2 forcing. We collect 2D surface and top-of-atmosphere fields, and 3D ocean temperature and salinity fields. Here, we document the collection of simulations and discuss initial results, including the evolution of surface and deep ocean temperature and cloud radiative effects. As of summer 2019, the collection includes 50 simulations of 15 models by 10 modeling centers. The data of LongRunMIP are publicly available. We encourage submission of more simulations in the future.

(Capsule Summary) LongRunMIP is the first collection of millennial-length simulations of complex coupled climate models and enables investigations of how these models equilibrate in response to radiative perturbations.

1. Motivation and objectives

Millennial-length climate simulations are necessary to understand the equilibrium states that occur in response to external forcings, as well as the relationship between transient and equilibrated behavior. Unforced millennial-length simulations are useful as well, as they allow us to consider long-term internal variability and to analyze shorter-term variability with increased statistical certainty. Reasons to study these long time scales include:

- To better understand long-term climate dynamics. Outstanding issues include the time scales of ocean circulation response (e.g., Jansen et al. 2018; Rind et al. 2018), continental drying trends (e.g., Sniderman et al. 2019), or sea level rise (e.g., Bilbao et al. 2015; Rugenstein et al. 2016c).

- To help predict the impacts of 20th and 21st century emissions on century timescales, such as ice sheet stability, deep ocean warming, or polar amplification (e.g., Frölicher and Joos 2010; Clark et al. 2016; Mauritsen and Pincus 2017), which are rarely explicitly simulated using a fully-coupled climate model.

- To more accurately estimate Equilibrium Climate Sensitivity (ECS), which is the equilibrium response of the surface air temperature to a doubling of CO_2 due to the "fast" feedbacks water vapor, lapse rate, clouds, and sea ice but excluding Earth system feedbacks such as changes in the carbon cycle, ice sheets, or vegetation. While ECS has long been a focus of scientific inquiry, substantial uncertainty remains as to its value (e.g., Charney et al. 1979; Knutti et al. 2017).

- To understand the relationship between the transient response of the climate and its equilibration. Since radiative feedbacks can depend on the evolution of the spatial pattern of warming (e.g., Senior and Mitchell 2000; Winton et al. 2010; Armour et al. 2013; Andrews et al. 2015; Andrews and Webb 2018) and on the background temperature (e.g., Colman and McAvaney 2009; Caballero and Huber 2013; Block and Mauritsen 2013; Meraner et al. 2013; Bloch-Johnson et al. 2015), a constant effective sensitivity of the climate is an inadequate assumption. Several methods have been proposed to predict the equilibrium response from transient simulations given a changing global feedback (Held et al. 2010; Winton et al. 2010; Armour et al. 2013; Geoffroy et al. 2013b,a; Frölicher et al. 2014; Proistosescu and Huybers 2017; Saint-Martin et al. 2019), but only fully equilibrated climate model simulations can serve to test how well these methods predict equilibrium conditions.

- To test theories for the relationship between feedbacks at different time-scales (Gregory et al. 2015, 2016; Zhou et al. 2016; Rugenstein et al. 2016a; Armour 2017; Proistosescu and Huybers 2017; Ceppi and Gregory 2017; Andrews and Webb 2018; Andrews et al. 2018), and to quantify the influence of slow, centennial-scale modes on the temperature evolution of the last century (Armour 2017; Proistosescu and Huybers 2017).

- To understand the relevance, time scales, and magnitude of the energy imbalances and drifts exhibited by climate models (e.g., Hobbs et al. 2016), with the potential application of decreasing the spin-up time needed to run these models.

- To understand the relationship between the forced response and internal variability. This relationship is currently studied using the time frame of one or two centuries, which is not enough to robustly quantify the internal variability under consideration (e.g., Maher et al. 2018; Lutsko and Takahashi 2018; Bloch-Johnson et al. in revision), millennial time scales with varying forcings (e.g., Köhler et al. 2017; Khon et al. 2018; Rehfeld et al. 2018) or by using expensive large ensemble simulations on decadal to centennial time scales (e.g., Deser et al. 2012; Maher et al. 2019; Rodgers et al. 2015). Millennial long simulations allow us to differentiate the transient response from the equilibrated forced response, even for quantities with large internal variability, such as precipitation, droughts, or the El Niño-Southern Oscillation (ENSO), and also the significance of a change in internal variability in a transient simulation relative to the control simulation (e.g., Brown et al. 2017).

- To compare climate model responses and paleo proxies, e.g. of surface or deep ocean temperatures or hydrological conditions on land in order to provide an independent way of testing climate models (Gebbie and Huybers 2019; Burls and Fedorov 2017; Scheff et al. 2017).

With LongRunMIP, we aim to advance knowledge in the above mentioned areas, fill a gap in the CMIP protocols (Taylor et al. 2011; Eyring et al. 2016), and collect published data in one location for easy public access.

The goals of LongRunMIP are

a) to continuously gather existing millennial-length simulations (both published and unpublished)

b) to standardize the collected data (e.g., using the same units and sign conventions)

c) to make the data publicly available and easily accessible

d) to foster an interdisciplinary community of users working on millennial-length problems, with experts on oceanography, atmospheric dynamics, energy balance modeling, ice sheet modeling, and paleoclimatology

The objectives of this paper are to

a) motivate the data collection strategy (Section 2)

b) specify the requirements for LongRunMIP contributors (Section 2 and b)

c) give an overview of currently submitted simulations and models (Section 2a, b, and Table 1)

d) give a sample of some initial analysis on these simulations (Section 3)

e) show how LongRunMIP relates to the existing literature on millennial-length simulations (Section 4a)

f) discuss the limitations and opportunities of LongRunMIP (Section 4b and c).

2. Experimental design and data collection strategy

LongRunMIP is the first and largest compilation of millennial-length simulations of complex climate models to date, where a "complex climate model" is understood to include an atmospheric, sea ice, land, and full depth ocean component, i.e. Atmosphere-Ocean General Circulation Models (AO-GCMs) with a dynamic atmosphere and ocean, as opposed to Models of Intermediate Complexity (EMICs), which are often used to study millennial-length questions in climate science (e.g., Zickfeld et al. 2013; Levermann et al. 2013). These model simulations include the "fast" feedbacks, such as changes in water vapor, lapse rate, sea ice, and clouds (Charney et al. 1979), but no "slow" feedbacks, such as changes in the ice-sheets. Vegetation is treated differently in the models (see Section 2b). In Section 4 we discuss the implications and limitations of our approach. Our goal is to collect as many simulations from as many independent models as possible, while keeping the archive and data sharing manageable. Consequently, we keep our requirements for contributions low.

a. Simulations and variables

A step-increase in atmospheric CO_2 concentrations (in the following called "step-forcing") is one of the simplest experiments for studying a model's response to forcing and is used as a benchmark simulation in CMIP3, CMIP5, and CMIP6 (Meehl et al. 2007; Taylor et al. 2011; Eyring et al. 2016). More realistic, gradual forcing scenarios have been shown to be representable by the step-forcing scenarios and exhibit feedbacks that correlate with those computed from step-forcing simulations (Good et al. 2013, 2015; Geoffroy and Saint-Martin 2014; Colman and Hanson 2016). The CMIP3 protocol required a step-forcing of *doubling* atmospheric CO_2 (here referred to as *abrupt2x*) above pre-industrial levels in a slab (i.e. non-dynamical) ocean, which for decades has been used to define ECS (e.g., Charney et al. 1979; Boer and Yu 2003c; Danabasoglu and Gent 2009). The integration time scale of these model setups are a couple of decades. However, a *quadrupling* of CO_2 (here referred to as *abrupt4x*) above pre-industrial levels has a better ratio of forced signal to internal variability. Because the forced response was assumed to scale linearly with increased forcing, the CMIP5 protocol requested an abrupt quadrupling of CO_2, now in a fully coupled model with a dynamical ocean, requiring longer integration time scales. The CMIP6 protocol again requests abrupt CO_2 quadrupling experiments, but encourages also the submission of abrupt CO_2 doublings, to study the relation between different forcing levels (Eyring et al. 2016; Good et al. 2016). CMIP5 and 6 protocols require the submission of 150 years of model output. A representative response of surface temperature anomalies and top of the atmosphere (TOA) radiative imbalance to an *abrupt4x* scenario is shown in Fig. 1. All anomalies mentioned in this paper are computed as the difference of the experiment from the average of the control simulation. After the 150 years of CMIP protocol length (blue shading) and after 1000 years (the minimum contribution to LongRunMIP, light red shading), the surface temperature response of the exemplary model shown here has reached 75 % and 88 % of its final value respectively, while the TOA radiation has equilibrated 85 % and 93 % of the forcing respectively (7.6 $W\,m^{-2}$ for this model). Thus, the final equilibration is a CPU-intensive exercise; the model shown here needs 4000 years to balance the final 0.5 $W\,m^{-2}$ (dark red shading).

The set of variables we collect is motivated by the interest of the LongRunMIP contributors and organizers in ECS, temperature and time dependent feedbacks, and deep ocean warming. Table 1 lists the variable names, units, and temporal and spatial resolution of the requested variables. The naming and sign conventions follow the CMIP5 protocol[1]. Given the large amount of data involved, we have kept our requested variable list low to allow as many groups as possible to participate. For the same reason, we do not request the data to be “CMORized”[2], i.e. written in conformance with the CMIP standards. However, we do homogenize signs, variable long names, and units, and also provide a regridded version of the fields, as well as global means.

b. Minimal, optimal, and current contributions

The *minimal requirement* to contribute to LongRunMIP are annual fields of a single simulation of any CO_2 forcing scenario that has at least 1000 years of constant forcing, along with a control simulation of any length. The complexity of the model should be CMIP5-class and include dynamic atmosphere, ocean, and sea ice components. An *optimal contribution* comprises monthly fields of fully equilibrated *abrupt2x*, *4x*, and *8x* simulations and a control simulation of several millennia.

Table 2 lists the model characteristics of the current contributions. Because the archive is assembled from experiments initiated independently for research purposes by multiple modeling groups, there is no pre-defined protocol like for the CMIP simulations. The models are diverse in origin

[1] http://cmip-pcmdi.llnl.gov/cmip5/data_description.html

[2] https://pcmdi.llnl.gov/CMIP6/Guide/dataUsers.html

and sample the CMIP5 range of models well (see discussion on model genealogy in Knutti 2010). Table 2 lists references for each model and publications using (parts of) the model output. Most of the current contributions to LongRunMIP are extensions of CMIP5 simulations, sometimes with updated model versions, while one model is an extension of a CMIP3 and another model an extension of a CMIP6 contributions (CCSM3 and CNRM-CM6-1 respectively).

Many of our current contributions fall short of the optimal expectation for equilibrium, because even several millennia are insufficient for the deep ocean to equilibrate (see discussion around Fig. 4). However, a few millennia appear to be enough for the surface temperature and TOA radiative imbalance to reach a new steady state in most models (see Section 3), and many questions can be adequately addressed with the current contributions. Our approach is to be inclusive, and to leave it to the user to determine the degree of equilibration needed for their research and to develop criteria for model selection.

Most contributions are step-forcing simulations, generally to 2x or 4x pre-industrial CO_2 concentrations (in Fig. 2 *abrupt2x* in colored in yellow, *abrupt4x* in orange, *abrupt8x* in dark red; *abrupt2.4x* and *abrupt4.8x* in dark and light pink). There are currently three exceptions: 1) some model simulations have gradual increases in CO_2 at 1% per year until doubled or quadrupled concentrations are reached, after which the concentration is kept constant (*1pct2x* and *1pct4x*, light and medium red in Fig. 2). 2) One model simulates the 1850-2010 period, after which CO_2 increases either piecewise linearly for 90 years until reaching 2.4x pre-industrial values (CCSM3II). 3) Finally, one model simulates the historical period and then the CMIP5 extended representative concentration pathway 8.5 (including CH_4, N_2O, CFC11, and CFC12 in addition to CO_2) until year 2300 after which all forcing agents are kept constant (*RCP8.5+*, violet in Fig. 2) For the models that did not contribute a a millennial-long step-forcing simulation, we collect short (typically 150 year) step-forcing simulations, generally from the CMIP5 archive. These simulations can be used to estimate the effective climate sensitivity and to relate transient and equilibrium responses. They are not mentioned in Table 2 and Fig. 2.

Most contributors were able to submit all requested variables. Some models only stored annual output, while for a few models the entire model output (including many more variables than listed in Table 1) is available. In principle, but with considerable effort, additional variables not listed in Table 1 could be requested from some or all contributors.

Some models are outliers in some sense. For example, the simulation *abrupt4x* of FAMOUS warms anomalously strong (Fig. 2 and 7) due to a shortwave cloud effect which is positive throughout the simulation and longwave clear-sky effect, which increases anomalously strongly (not shown, see Rugenstein et al. (2019)). In principle though, such extreme behavior could represent possible characteristics of the real world (e.g., Bloch-Johnson et al. 2015; Schneider et al. 2019). Another atypical model is EC-Earth-PISM, which is the only model with an interactive Greenland ice sheet. This additional component and its historical and RCP8.5+ forcing scenario makes it harder to compare the simulation to other models and attribute changes to one forcing component. This model also does not equilibrate but finally produces a negative TOA imbalance, which probably would increase if the simulation was integrated further. We encourage similar "problematic" submissions, since our focus is on understanding model behavior and the large range of model responses (discussed in Section 3).

In nine models, the vegetation is fixed to pre-industrial conditions (ECHAM5, CCSM3, CCSM3II, HadCM3L, FAMOUS, MIROC32, ECEARTH, GISSE2R, CNRMCM61), while the other seven models have dynamic vegetation schemes (MPIESM11, MPIESM12, CESM104, HadGEM2, GFDLESM2M, GFDLCM3, IPSLCM5A).

3. Sample of model output

a. Imbalances in the control simulation and drift

In principle, the TOA radiative imbalance should be zero in a control simulation. Most models contributing to LongRunMIP do not loose or gain energy (Fig. 3). However, some models that are equilibrated in the sense that they show no substantial drift, still have a constant energy leakage. For CMIP5 models, imbalances of the same order of magnitude (and larger) have been shown to be uncorrelated with the forced response (Hobbs et al. 2016). If computing atmospheric anomalies, we suggest users to take the difference of each time step to the time-averaged control simulation imbalance, except for CCSM3II and GFDL-CM3 for which the difference to a polynomial fit to the control simulation time series seems appropriate (see Fig. 3).

The deep ocean (defined here as depth level around 2 km) has an astonishingly small drift in the global average in most models (Fig. 4, lowest panel). While the surface ocean time scales closely follows the global mean surface air temperature anomaly, the deep ocean takes centuries to equilibrate. Panel a and b of Fig. 4 display the surface and deep ocean temperature anomalies, computed as the difference of the forced and control simulations, while the lowest panel shows the absolute temperatures of the deep ocean in the control simulations to indicate the model spread in the base state. Previous work on long-term trends in deep ocean temperature and salinity shows that these trends may reflect ongoing changes in stratification and the strength and depth of the Atlantic Meridional Overturning Circulation (AMOC; e.g., Stouffer and Manabe 2003; Rugenstein et al. 2016a; Marzocchi and Jansen 2017; Jansen et al. 2018). Even if the energy flux imbalance at the TOA or the ocean surface are close to a new steady state this does not necessarily indicate that the deep ocean is equilibrated as well (Zhang et al. 2013; Hobbs et al. 2016; Marzocchi and Jansen 2017). Reaching deep ocean equilibration may not be necessary for studies concerned with surface properties only. However, for interpretation of paleo proxies and comparison with model simulations, distinguishing between the transient and equilibrium response in the intermediate or deep ocean is necessary (Zhang et al. 2013; Marzocchi and Jansen 2017; Rind et al. 2018; Jansen et al. 2018).

b. Evolution of surface temperature and cloud radiative effect

The evolution of large scale surface air temperature patterns on decadal to millennial time scales (Fig. 5) are robust among models and different forcing levels. The simulations show a strong land-sea warming contrast on short time scales and little warming over the Southern Ocean on decadal to centennial time scales (e.g., Manabe et al. 1991; Gregory 2000; Joshi and Gregory 2008; Geoffroy and Saint-Martin 2014; Armour et al. 2016). A warming pattern reminiscent of the positive phase of ENSO and the Interdecadal Pacific Oscillation occurs throughout the Pacific basin (panel b; Held et al. 2010; Song and Zhang 2014; Andrews et al. 2015; Luo et al. 2017) but decays on centennial to millennial time scales (panel c and d), with a large model spread in time scales (not shown). As it approaches equilibrium, the temperature pattern becomes more homogeneous, the land-sea warming contrast reduces (e.g., Held et al. 2010; Geoffroy and Saint-Martin 2014), and the Southern Hemisphere high latitudes keep warming beyond year 1000. As in previous studies, the AMOC first declines (Gregory et al. 2005; Zhu et al. 2014; Kostov et al. 2014; Trossman et al. 2016) and then recovers (Stouffer and Manabe 2003; Li et al. 2013; Zickfeld et al. 2013; Rugenstein et al. 2016a; Rind et al. 2018), resulting in a delayed warming in the North Atlantic. Panel a, b, and e correspond to the blue shading in Fig. 1, and are known from CMIP5 simulations (e.g., Andrews et al. 2015), while panel c, d, f, and g highlight that the simulations still warm substantially on centennial to millennial time scales, mainly in areas with more sensitive – i.e. positive or small negative – feedbacks (Rugenstein et al. 2019). Normalizing the zonal-mean temperature anomaly by the global mean warming reveals the relative zonal-mean warming (Fig. 6). Arctic amplification begins very early in the simulations and warming throughout the Southern Hemisphere is lower than the global average in almost all models for the first centuries. Between year 100 and 1000 the Southern Hemisphere warms more than the Northern Hemisphere in all latitudes poleward from 30°, in some regions by more than 4 K. Antarctic warming slowly increases, but is still substantially less than Arctic amplification (e.g., Salzmann 2017). In a couple of models, the amplitude of Antarctic and Arctic amplification is the same after 4000 years of model integration time (GISSE2R and ECHAM5; Li et al. 2013), while in other models the Antarctic amplification stays substantially smaller and still increasing after a couple of thousand years. LongRunMIP shows that there is no reduction in model spread in the polar regions through time and that although all models follow a similar large scale pattern evolution (Fig. 5), the local response time scales, e.g. in the North Atlantic, Southern Ocean, or equatorial Pacific differ by hundreds to thousands years.

While the large scale temperature response is rather robust between models and simulations, the cloud radiative effect (CRE) differs strongly in magnitude and time evolution, both between models and between forcing levels for the same model (Fig 7). We show the shortwave CRE – computed as the difference between "all sky" and "clear sky" shortwave radiative fluxes (e.g., Ramanathan et al. 1989; Ceppi et al. 2017) – as a function of surface air temperature anomaly. The models disagree in the overall sign, as expected from CMIP5 models on shorter time scales (e.g., Vial et al. 2013; Caldwell et al. 2015), but can even change sign within a single simulation (e.g., ECEARTH or CESM *abrupt8x*). The strength of variation in time within one simulation can depend strongly on the forcing level (e.g. MIROC32 *1pct2x* vs. *1pct4x*) and the time scales of change differ between the models (e.g. IPSLCM5R vs. MPIESM12 *abrupt4x*). For some simulations, cloud response barely changes with temperature, contributing negligibly to the overall feedback (e.g. MPIESM12 *abrupt16x*, CESM104 *abrupt4x*, and MIROC32 *1pct2x*).

4. Discussion and Outlook

a. Published millennial-length simulations

Models of intermediate complexity are the most common tools used to study century to millennium time scales in the climate system (e.g., Zickfeld et al. 2013; Eby et al. 2013; Levermann et al. 2013; Rugenstein et al. 2016c; Jansen et al. 2018). However, they usually have a poorly resolved atmosphere and little or no representation of cloud processes. In contrast, the publications in Table 3 feature millennium-length AO-GCM simulations. Asterisks mark contributions to LongRunMIP. These papers provide a solid body of work on millennial-length climate simulations, but rarely use the same forcing levels and simulation length and focus on different aspects of the climate system. Three papers compare model formulation and processes of two AO-GCMs each (Frölicher et al. 2014; Paynter et al. 2018; Krasting et al. 2018), but otherwise models have not been systematically compared against each other. Fig. 4 and 7 show that AO-GCMs can strongly differ in their behavior. Spatial patterns of e.g., precipitation and surface heat fluxes also vary strongly between models and between different forcing scenarios for the same model (not shown), suggesting that some mechanisms and processes discussed in the published literature are not generalizable across models. For example, there is disagreement about which regions are thought to dominate the changing feedback parameter (Senior and Mitchell 2000; Andrews et al. 2015; Meraner et al. 2013; Caballero and Huber 2013) or whether or not, and on which time scales, the AMOC recovers from its initial reduction (Voss and Mikolajewicz 2001; Stouffer and Manabe 2003; Li et al. 2013; Rind et al. 2018; Thomas and Fedorov 2019). Paleo climate simulations are often several thousand years long, however, they usually include boundary conditions such as ice sheets or changing continental configurations, which differ from the ones used here. However, paleo climate studies often discuss equilibration time scales and deep ocean temperature trends relevant to the types of models included in LongRunMIP (e.g., Brandefelt and Otto-Bliesner 2009; Zhang et al. 2013; Klockmann et al. 2016; Marzocchi and Jansen 2017; Gottschalk et al. 2019).

b. Limitations

LongRunMIP analyses are currently limited mainly by the collected *variables* (Table 1). Including cloud fields and 3D atmospheric temperature and humidity fields, for example, would allow users to study atmospheric dynamics and radiative feedbacks in more detail. The *different forcing scenarios* of model contributions to LongRunMIP are both a strength and weakness. Minimal requirements have encouraged a large number of contributions so far. However, studying a single forcing scenario requires model selection or scaling between different forcing levels. *Slab ocean simulations*, which replace a model's dynamical ocean with a much shallower non-dynamical mixed-layer, are a computationally cheap tool to compare fast and slow time scales and the relevance of surface warming patterns (Boer and Yu 2003c; Danabasoglu and Gent 2009; Li et al. 2013). We hope to receive submissions of these simulations in the future, to allow analysis of their utility. Century to millennial-time scales in the real world include more processes and *Earth System Feedbacks* than are included in LongRunMIP simulations, such as the carbon cycle, vegetation feedbacks, forcing agents other than CO_2 (such as other greenhouse gases or aerosols), ice sheets, glacial rebound effects, changes to continental configuration, and orbital variation. Further, the real climate system is never in equilibrium or steady state, because the forcing continuously changes (e.g., Köhler et al. 2017). These Earth system feedbacks and additional forcings must be taken into account when comparing the LongRunMIP models with paleo proxies or when projecting or predicting changes in future centuries or millennia.

c. Summary and expected impact

LongRunMIP is the first archive of millennial-length simulations of complex climate models, featuring 50 simulations of 15 models by 10 modeling centers under various forcing scenarios (Table 2). The archive provides an unprecedented opportunity to study the equilibrium response of a large number of models to forcing. The variables included allow study of a range of phenomena associated with the atmosphere, ocean, land, and sea ice (Table 1), and we expect LongRunMIP to contribute to current discussions laid out in Section 1. This includes ocean heat uptake, sea level rise, ocean circulation response to warming, large scale modes of variability, sea ice reduction, polar amplification, precipitation variability, atmospheric dynamics, long-term memory in time series, spatial warming patterns, ocean - atmosphere interactions, model spin-up techniques, the relation of internal variability and forced response under different forcing levels, committed climate response, and the relation of time and state dependence of fast feedbacks and Earth System Feedbacks and processes.

LongRunMIP is a MIP of opportunity, without an argeed upon protocol, and is a result of the willingness of individual research groups to provide model output from simulations often conducted over years of real-world time. As a result, the experiments are not standardized, but most models provided a millennial-length simulation that begins with an abrupt quadrupling of CO_2 concentration. In addition to collecting simulations, we provide output with standardized formats and variable names, and include versions regridded to a common grid, as well as global averages.

LongRunMIP builds upon a body of pioneering studies that looked at the behavior of models beyond the centennial scale (Table 3), LongRunMIP allows this sort of analysis to be applied across a diverse group of models that exhibit strikingly different behavior (Fig. 7), and hopefully encourage others to look beyond the limitations and assumptions normally imposed by computational constraints, to directly study the equilibration of the fully coupled atmosphere-ocean system.

Data access and sharing

LongRunMIP currently consists of 15 TB of data and available for download at https://data.iac.ethz.ch/longrunmip/. Fields shown in this paper can be accessed on https://data.iac.ethz.ch/longrunmip/BAMS/.

See www.longrunmip.org for more details on available variables, contact information, sample figures and videos, and links to join a discussion community. We will be collecting more simulations over the next couple of years.

Acknowledgments. MR is funded by the Alexander von Humboldt Foundation. TA was supported by the Joint UK BEIS/Defra Met Office Hadley Centre Climate Programme (GA01101). NCAR is a major facility sponsored by the US National Science Foundation under Cooperative Agreement No. 1852977 TLF acknowledges support from the Swiss National Science Foundation under grant PP00P2_170687, from the EU-H2020 project CCiCC, and from the Swiss National Supercomputing Centre (CSCS). CL was supported through the Clusters of Excellence CliSAP (EXC177) and CLICCS (EXC2037), University Hamburg, funded through the German Research Foundation (DFG). SY was partly supported by European Research Council under the European Community's Seventh Framework Programme (FP7/20072013)/ERC grant agreement 610055 as part of the ice2ice project. This work was made possible for IPSL thanks to the HPC resources of TGCC and IDRIS made available by GENCI (Grand Equipement National de Calcul Intensif), CEA (Commissariat à l'Energie Atomique et aux Energies Alternatives) and CNRS (Centre National de la Recherche Scientifique) (project 016178).

References

Andrews, T., J. M. Gregory, and M. J. Webb, 2015: The dependence of radiative forcing and feedback on evolving patterns of surface temperature change in climate models. *Journal of Climate*, **28 (4)**, 1630–1648, URL http://dx.doi.org/10.1175/JCLI-D-14-00545.1.

Andrews, T., and M. J. Webb, 2018: The Dependence of Global Cloud and Lapse Rate Feedbacks on the Spatial Structure of Tropical Pacific Warming. *Journal of Climate*, **31 (2)**, 641–654, doi:10.1175/JCLI-D-17-0087.1, URL https://doi.org/10.1175/JCLI-D-17-0087.1.

Andrews, T., and Coauthors, 2018: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity. *Geophysical Research Letters*, **45 (16)**, 8490–8499, doi:10.1029/2018GL078887, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078887.

Armour, K. C., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate feedbacks. *Nature Climate Change*, **7**, 331 EP –, URL http://dx.doi.org/10.1038/nclimate3278.

Armour, K. C., C. M. Bitz, and G. H. Roe, 2013: Time-Varying Climate Sensitivity from Regional Feedbacks. *Journal of Climate*, **26 (13)**, 4518–4534, URL http://dx.doi.org/10.1175/JCLI-D-12-00544.1.

Armour, K. C., J. Marshall, J. R. Scott, A. Donohoe, and E. R. Newsom, 2016: Southern ocean warming delayed by circumpolar upwelling and equatorward transport. *Nature Geosci*, **9 (7)**, 549–554, URL http://dx.doi.org/10.1038/ngeo2731.

Bi, D., W. F. Budd, A. C. Hirst, and X. Wu, 2001: Collapse and reorganisation of the Southern Ocean overturning under global warming in a coupled model. *Geophysical Research Letters*, **28 (20)**, 3927–3930, URL http://dx.doi.org/10.1029/2001GL013705.

Bilbao, R. A., J. M. Gregory, and N. Bouttes, 2015: Analysis of the regional pattern of sea level change due to ocean dynamics and density change for 1993–2099 in observations and CMIP5 AOGCMs. *Climate Dynamics*, **45 (9-10)**, 2647–2666, URL http://dx.doi.org/10.1007/s00382-015-2499-z.

Bloch-Johnson, J., R. T. Pierrehumbert, and D. S. Abbot, 2015: Feedback temperature dependence determines the risk of high warming. *Geophysical Research Letters*, **42 (12)**, 4973– 4980, doi: 10.1002/2015GL064240, URL http://dx.doi.org/10.1002/2015GL064240, 2015GL064240.

Bloch-Johnson, J., M. Rugenstein, and D. S. Abbot, in revision: Spatial radiative feedbacks from interannual variability using multiple regression. *Journal of Climate*.

Block, K., and T. Mauritsen, 2013: Forcing and feedback in the MPI-ESM-LR coupled model under abruptly quadrupled CO_2. *Journal of Advances in Modeling Earth Systems*, **5 (4)**, 676–691, URL http://dx.doi.org/10.1002/jame.20041.

Boer, G., and B. Yu, 2003a: Climate sensitivity and climate state. *Climate Dynamics*, **21 (2)**, 167–176, URL http://dx.doi.org/10.1007/s00382-003-0323-7.

Boer, G., and B. Yu, 2003b: Climate sensitivity and response. *Climate Dynamics*, **20 (4)**, 415–429, URL http://dx.doi.org/10.1007/s00382-002-0283-3.

Boer, G. J., and B. Yu, 2003c: Dynamical aspects of climate sensitivity. *Geophysical Research Letters*, **30 (3)**, doi:10.1029/2002GL016549, URL http://dx.doi.org/10.1029/2002GL016549.

Brandefelt, J., and B. L. Otto-Bliesner, 2009: Equilibration and variability in a last glacial maximum climate simulation with ccsm3. *Geophysical Research Letters*, **36 (19)**, doi:10.1029/2009GL040364, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009GL040364.

Brown, P. T., Y. Ming, W. Li, and S. A. Hill, 2017: Change in the magnitude and mechanisms of global temperature variability with warming. *Nature Climate Change*, **7**, 743 EP –, URL http://dx.doi.org/10.1038/nclimate3381.

Burls, N. J., and A. V. Fedorov, 2017: Wetter subtropics in a warmer world: Contrasting past and future hydrological cycles. *Proceedings of the National Academy of Sciences*, **114 (49)**, 12 888–12 893, doi:10.1073/pnas.1703421114, URL https://www.pnas.org/content/114/49/12888.

Caballero, R., and M. Huber, 2013: State-dependent climate sensitivity in past warm climates and its implications for future climate projections. *Proceedings of the National Academy of Sciences of the United States of America*, **110 (35)**, 14 162–14 167, URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3761583/.

Caldwell, P. M., M. D. Zelinka, K. E. Taylor, and K. Marvel, 2015: Quantifying the sources of inter-model spread in equilibrium climate sensitivity. *Journal of Climate*, doi:10.1175/JCLI-D-15-0352.1, URL http://dx.doi.org/10.1175/JCLI-D-15-0352.1.

Cao, L., L. Duan, G. Bala, and K. Caldeira, 2016: Simulated long-term climate response to idealized solar geoengineering. *Geophysical Research Letters*, URL http://dx.doi.org/10.1002/2016GL068079.

Castruccio, S., D. J. McInerney, M. L. Stein, F. Liu Crouch, R. L. Jacob, and E. J. Moyer, 2014: Statistical Emulation of Climate Model Projections Based on Precomputed GCM Runs. *Journal* *of Climate*, **27 (5)**, 1829–1844, doi:10.1175/JCLI-D-13-00099.1, URL https://doi.org/10.1175/JCLI-D-13-00099.1, https://doi.org/10.1175/JCLI-D-13-00099.1.

Ceppi, P., F. Brient, M. D. Zelinka, and D. L. Hartmann, 2017: Cloud feedback mechanisms and their representation in global climate models. *Wiley Interdisciplinary Reviews: Climate Change*, **8 (4)**, e465, doi:10.1002/wcc.465, URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wcc.465, https://onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.465.

Ceppi, P., and J. M. Gregory, 2017: Relationship of tropospheric stability to climate sensitivity and earth's observed radiation budget. *Proceedings of the National Academy of Sciences*, **114 (50)**, 13 126–13 131, doi:10.1073/pnas.1714308114, URL https://www.pnas.org/content/114/50/13126.

Charney, J., and Coauthors, 1979: Carbon Dioxide and Climate: A Scientific Assessment. Tech. rep., National Academy of Science, Washington, DC.

Clark, P. U., and Coauthors, 2016: Consequences of twenty-first-century policy for multi-millennial climate and sea-level change. *Nature Climate Change*, **6**, 360 EP –, URL http://dx.doi.org/10.1038/nclimate2923.

Collins, W. J., and Coauthors, 2011: Development and evaluation of an Earth-System model – HadGEM2. *Geoscientific Model Development*, **4 (4)**, 1051–1075, URL https://www.geosci-model-dev.net/4/1051/2011/.

Colman, R., and L. Hanson, 2016: On the relative strength of radiative feedbacks under climate variability and change. *Climate Dynamics*, 1–15, doi:10.1007/s00382-016-3441-8, URL http://dx.doi.org/10.1007/s00382-016-3441-8.

Colman, R., and B. McAvaney, 2009: Climate feedbacks under a very broad range of forcing. *Geophysical Research Letters*, **36 (1)**, n/a–n/a, doi:10.1029/2008GL036268, URL http://dx.doi.org/10.1029/2008GL036268.

Cox, P. M., R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell, 2000: Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, **408**, 184 EP –, URL http://dx.doi.org/10.1038/35041539.

Danabasoglu, G., and P. R. Gent, 2009: Equilibrium Climate Sensitivity: Is It Accurate to Use a Slab Ocean Model? *Journal of Climate*, **22 (9)**, 2494–2499, URL http://dx.doi.org/10.1175/2008JCLI2596.1.

Danabasoglu, G., S. G. Yeager, Y.-O. Kwon, J. J. Tribbia, A. S. Phillips, and J. W. Hurrell, 2012: Variability of the Atlantic Meridional Overturning Circulation in CCSM4. *Journal of Climate*, URL http://dx.doi.org/10.1175/JCLI-D-11-00463.1.

Deser, C., A. Phillips, V. Bourdette, and H. Teng, 2012: Uncertainty in climate change projections: the role of internal variability. *Climate Dynamics*, **38 (3)**, 527–546, doi:10.1007/s00382-010-0977-x, URL https://doi.org/10.1007/s00382-010-0977-x.

Donner, L. J., and Coauthors, 2011: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate*, **24 (13)**, 3484–3519, URL https://doi.org/10.1175/2011JCLI3955.1.

Dufresne, J.-L., and Coauthors, 2013: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, **40 (9)**, 2123–2165, URL https://doi.org/10.1007/s00382-012-1636-1.

Dunne, J. P., and Coauthors, 2012: GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics. *Journal of Climate*.

Eby, M., and Coauthors, 2013: Historical and idealized climate model experiments: an intercomparison of Earth system models of intermediate complexity. *Climate of the Past*, **9 (3)**, 1111–1140, URL http://www.clim-past.net/9/1111/2013/.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9 (5)**, 1937–1958, doi:10.5194/gmd-9-1937-2016, URL https://www.geosci-model-dev.net/9/1937/2016/.

Frölicher, T., and F. Joos, 2010: Reversible and irreversible impacts of greenhouse gas emissions in multi-century projections with the NCAR global coupled carbon cycle-climate model. *Climate Dynamics*, **35 (7-8)**, 1439–1459, doi:10.1007/s00382-009-0727-0, URL http://dx.doi.org/10.1007/s00382-009-0727-0.

Frölicher, T. L., and D. J. Paynter, 2015: Extending the relationship between global warming and cumulative carbon emissions to multi-millennial timescales. *Environmental Research Letters*, **10 (7)**, 075 002, URL http://stacks.iop.org/1748-9326/10/i=7/a=075002.

Frölicher, T. L., M. Winton, and J. L. Sarmiento, 2014: Continued global warming after CO_2 emissions stoppage. *Nature Clim. Change*, **4 (1)**, 40–44, URL http://dx.doi.org/10.1038/nclimate2060.

Gebbie, G., and P. Huybers, 2019: The Little Ice Age and 20th-century deep Pacific cooling. *Science*, **363 (6422)**, 70–74, doi:10.1126/science.aar8413, URL http://science.sciencemag.org/content/363/6422/70.

Gent, P. R., and Coauthors, 2011: The Community Climate System Model Version 4. *Journal of Climate*, **24 (19)**, 4973–4991, URL http://dx.doi.org/10.1175/2011JCLI4083.1.

Geoffroy, O., and D. Saint-Martin, 2014: Pattern decomposition of the transient climate response. *Tellus A: Dynamic Meteorology and Oceanography*, **66 (1)**, 23 393, doi:10.3402/tellusa.v66.23393, URL https://doi.org/10.3402/tellusa.v66.23393.

Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. Olivié, and S. Tytéca, 2013a: Transient Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Efficacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, **26 (6)**, 1859–1876, URL http://dx.doi.org/10.1175/JCLI-D-12-00196.1.

Geoffroy, O., D. Saint-Martin, D. J. L. Olivié, A. Voldoire, G. Bellon, and S. Tytéca, 2013b: Transient Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Parameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, **26 (6)**, 1841–1857, URL http://dx.doi.org/10.1175/JCLI-D-12-00195.1.

Gillett, N. P., V. K. Arora, K. Zickfeld, S. J. Marshall, and W. J. Merryfield, 2011: Ongoing climate change following a complete cessation of carbon dioxide emissions. *Nature Geosci*, **4 (2)**, 83–87, URL http://dx.doi.org/10.1038/ngeo1047.

Good, P., T. Andrews, R. Chadwick, J.-L. Dufresne, J. M. Gregory, J. A. Lowe, N. Schaller, and H. Shiogama, 2016: nonlinMIP contribution to CMIP6: model intercomparison project for non-linear mechanisms: physical basis, experimental design and analysis principles (v1.0). *Geoscientific Model Development*, **9 (11)**, 4019–4028, doi:10.5194/gmd-9-4019-2016, URL https://www.geosci-model-dev.net/9/4019/2016/.

Good, P., J. Gregory, J. Lowe, and T. Andrews, 2013: Abrupt CO_2 experiments as tools for predicting and understanding CMIP5 representative concentration pathway projections. *Climate Dynamics*, **40 (3-4)**, 1041–1053, URL http://dx.doi.org/10.1007/s00382-012-1410-4.

Good, P., and Coauthors, 2015: Nonlinear regional warming with increasing co2 concentrations. *Nature Clim. Change*, **5 (2)**, 138–142, URL http://dx.doi.org/10.1038/nclimate2498.

Gottschalk, J., and Coauthors, 2019: Mechanisms of millennial-scale atmospheric co2 change in numerical model simulations. *Quaternary Science Reviews*, **220**, 30 – 74, doi:https://doi.org/10.1016/j.quascirev.2019.05.013, URL http://www.sciencedirect.com/science/article/pii/S0277379118310473.

Gregory, J. M., 2000: Vertical heat transports in the ocean and their effect on time-dependent climate change. *Climate Dynamics*, **16**, 501–515, URL http://dx.doi.org/10.1007/s003820000059, 10.1007/s003820000059.

Gregory, J. M., T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response parameter under increasing CO_2. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, **373 (2054)**, URL http://rsta.royalsocietypublishing.org/content/373/2054/20140417.

Gregory, J. M., T. Andrews, P. Good, T. Mauritsen, and P. M. Forster, 2016: Small global-mean cooling due to volcanic radiative forcing. *Climate Dynamics*, 1–13, URL http://dx.doi.org/10.1007/s00382-016-3055-1.

Gregory, J. M., and Coauthors, 2004: A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, **31 (3)**, URL http://dx.doi.org/10.1029/2003GL018747.

Gregory, J. M., and Coauthors, 2005: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentration. *Geophys. Res. Lett.*, **32 (L12703)**.

Hasumi, H., and S. Emori, 2004: K-1 Coupled GCM (MIROC) Description. Tech. rep., Center for Climate System Research CCSR, University of Tokyo, National Institute for Envoronmeltal Studies (NIES), Frontier Research Center for Global Change (FRCGC).

Hazeleger, W., and Coauthors, 2012: EC-Earth V2.2: description and validation of a new seamless earth system prediction model. *Climate Dynamics*, **39 (11)**, 2611–2629, URL https://doi.org/10.1007/s00382-011-1228-5.

Held, I., M. Winton, K. Takahashi, T. L. Delworth, F. Zeng, and G. Vallis, 2010: Probing the Fast and Slow Components of Global Warming by Returning Abruptly to Preindustrial Forcing. *Journal of Climate*, **23**, 2418 – 2427.

Hobbs, W., M. D. Palmer, and D. Monselesan, 2016: An Energy Conservation Analysis of Ocean Drift in the CMIP5 Global Coupled Models. *Journal of Climate*, **29 (5)**, 1639–1653, doi:10.1175/JCLI-D-15-0477.1, URL http://dx.doi.org/10.1175/JCLI-D-15-0477.1.

Jansen, M. F., L.-P. Nadeau, and T. M. Merlis, 2018: Transient versus Equilibrium Response of the Ocean's Overturning Circulation to Warming. *Journal of Climate*, **31 (13)**, 5147–5163, doi:10.1175/JCLI-D-17-0797.1, URL https://doi.org/10.1175/JCLI-D-17-0797.1.

Jonko, A. K., K. M. Shell, B. M. Sanderson, and G. Danabasoglu, 2013: Climate Feedbacks in CCSM3 under Changing CO2 Forcing. Part II: Variation of Climate Feedbacks and Sensitivity with Forcing. *Journal of Climate*, **26 (9)**, 2784–2795, URL http://dx.doi.org/10.1175/JCLI-D-12-00479.1.

Joshi, M., and J. Gregory, 2008: Dependence of the land-sea contrast in surface climate response on the nature of the forcing. *Geophysical Research Letters*, **35 (24)**, doi:10.1029/2008GL036234, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2008GL036234.

Jungclaus, J. H., and Coauthors, 2006: Ocean Circulation and Tropical Variability in the Coupled Model ECHAM5/MPI-OM. *Journal of Climate*, **19 (16)**, 3952–3972, URL https://doi.org/10.1175/JCLI3827.1.

Khon, V. C., B. Schneider, M. Latif, W. Park, and C. Wengel, 2018: Evolution of Eastern Equatorial Pacific Seasonal and Interannual Variability in Response to Orbital Forcing During the Holocene and Eemian From Model Simulations. *Geophysical Research Letters*, **45 (18)**, 9843–9851, doi:10.1029/2018GL079337, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL079337.

Klockmann, M., U. Mikolajewicz, and J. Marotzke, 2016: The effect of greenhouse gas concentrations and ice sheets on the glacial AMOC in a coupled climate model. *Climate of the Past*, **12 (9)**, 1829–1846, doi:10.5194/cp-12-1829-2016, URL https://www.clim-past.net/12/1829/2016/.

Knutti, R., 2010: The end of model democracy? *Climatic Change*, **102**, 395–404.

Knutti, R., M. A. A. Rugenstein, and G. C. Hegerl, 2017: Beyond equilibrium climate sensitivity. *Nature Geoscience*, **10**, 727 EP –, URL http://dx.doi.org/10.1038/ngeo3017.

Köhler, P., C. Nehrbass-Ahles, J. Schmitt, T. F. Stocker, and H. Fischer, 2017: A 156 kyr smoothed history of the atmospheric greenhouse gases CO_2, CH_4, and N_2O and their radiative forcing. *Earth System Science Data*, **9 (1)**, 363–387, doi:10.5194/essd-9-363-2017, URL https://www.earth-syst-sci-data.net/9/363/2017/.

Kostov, Y., K. C. Armour, and J. Marshall, 2014: Impact of the atlantic meridional overturning circulation on ocean heat storage and transient climate change. *Geophysical Research Letters*, **41 (6)**, 2108–2116, doi:10.1002/2013GL058998, URL http://dx.doi.org/10.1002/2013GL058998.

Krasting, J. P., R. J. Stouffer, S. M. Griffies, R. W. Hallberg, S. L. Malyshev, B. L. Samuels, and L. T. Sentman, 2018: Role of Ocean Model Formulation in Climate Response Uncertainty. *Journal of Climate*, **31 (22)**, 9313–9333, doi:10.1175/JCLI-D-18-0035.1, URL https://doi.org/10.1175/JCLI-D-18-0035.1, https://doi.org/10.1175/JCLI-D-18-0035.1.

Levermann, A., P. U. Clark, B. Marzeion, G. A. Milne, D. Pollard, V. Radic, and A. Robinson, 2013: The multimillennial sea-level commitment of global warming. *Proceedings of the National Academy of Sciences*, **110 (34)**, 13 745–13 750, URL http://www.pnas.org/content/110/34/13745.abstract.

Li, C., J.-S. Storch, and J. Marotzke, 2013: Deep-ocean heat uptake and equilibrium climate response. *Climate Dynamics*, **40 (5-6)**, 1071–1086, URL http://dx.doi.org/10.1007/s00382-012-1350-z.

Luo, Y., J. Lu, F. Liu, and O. Garuba, 2017: The Role of Ocean Dynamical Thermostat in Delaying the El Niño–Like Response over the Equatorial Pacific to Climate Warming. *Journal of Climate*, **30 (8)**, 2811–2827, doi:10.1175/JCLI-D-16-0454.1, URL https://doi.org/10.1175/JCLI-D-16-0454.1.

Lutsko, N. J., and K. Takahashi, 2018: What Can the Internal Variability of CMIP5 Models Tell Us about Their Climate Sensitivity? *Journal of Climate*, **31 (13)**, 5051–5069, doi:10.1175/JCLI-D-17-0736.1, URL https://doi.org/10.1175/JCLI-D-17-0736.1.

Maher, N., D. Matei, S. Milinski, and J. Marotzke, 2018: Enso change in climate projections: Forced response or internal variability? *Geophysical Research Letters*, **45 (20)**, 11,390–11,398, doi:10.1029/2018GL079764, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL079764.

Maher, N., and Coauthors, 2019: The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability. *Journal of Advances in Modeling Earth Systems*, **0 (0)**, doi:10.1029/2019MS001639, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001639.

Manabe, S., R. J. Stouffer, M. J. Spelman, and K. Bryan, 1991: Transient Responses of a Coupled Ocean Atmosphere Model to Gradual Changes of Atmospheric CO_2. Part I. Annual Mean Response. *Journal of Climate*, **4 (8)**, 785–818.

Marzocchi, A., and M. F. Jansen, 2017: Connecting antarctic sea ice to deep-ocean circulation in modern and glacial climate simulations. *Geophysical Research Letters*, **44 (12)**, 6286–6295, doi:10.1002/2017GL073936, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017GL073936.

Mauritsen, T., and R. Pincus, 2017: Committed warming inferred from observations. *Nature Climate Change*, **7**, 652 EP –, URL https://doi.org/10.1038/nclimate3357.

Mauritsen, T., and Coauthors, 2018: Developments in the mpi-m earth system model version 1.2 (mpi-esm 1.2) and its response to increasing co2. *Journal of Advances in Modeling Earth Systems*, **0 (ja)**, doi:10.1029/2018MS001400, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001400.

Meehl, G. A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J. F. B. Mitchell, R. J. Stouffer, and K. E. Taylor, 2007: THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bulletin of the American Meteorological Society*, **88 (9)**, 1383–1394, doi:10.1175/BAMS-88-9-1383, URL https://doi.org/10.1175/BAMS-88-9-1383.

Meraner, K., T. Mauritsen, and A. Voigt, 2013: Robust increase in equilibrium climate sensitivity under global warming. *Geophysical Research Letters*, **40 (22)**, 5944–5948, URL http://dx.doi.org/10.1002/2013GL058118.

Miller, R. L., and Coauthors, 2014: CMIP5 historical simulations (1850–2012) with GISS ModelE2. *Journal of Advances in Modeling Earth Systems*, **6 (2)**, 441–478, URL http://dx.doi.org/10.1002/2013MS000266.

Nazarenko, L., and Coauthors, 2015: Future climate change under RCP emission scenarios with GISS ModelE2. *Journal of Advances in Modeling Earth Systems*, **7 (1)**, 244–267, URL http://dx.doi.org/10.1002/2014MS000403.

Paynter, D., T. L. Frölicher, L. W. Horowitz, and L. G. Silvers, 2018: Equilibrium Climate Sensitivity Obtained From Multimillennial Runs of Two GFDL Climate Models. *Journal of Geophysical Research: Atmospheres*, **123 (4)**, 1921–1941, doi:10.1002/2017JD027885, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017JD027885.

Proistosescu, C., and P. J. Huybers, 2017: Slow climate mode reconciles historical and model-based estimates of climate sensitivity. *Science Advances*, **3 (7)**, doi:10.1126/sciadv.1602821, URL http://advances.sciencemag.org/content/3/7/e1602821.

Ramanathan, V., R. D. Cess, E. F. Harrison, P. Minnis, B. R. Barkstrom, E. Ahmad, and D. Hartmann, 1989: Cloud-Radiative Forcing and Climate: Results from the Earth

Radiation Budget Experiment. *Science*, **243 (4887)**, 57–63, doi:10.1126/science.243.4887.57, URL https://science.sciencemag.org/content/243/4887/57, https://science.sciencemag.org/content/243/4887/57.full.pdf.

Rehfeld, K., T. Münch, S. L. Ho, and T. Laepple, 2018: Global patterns of declining temperature variability from the Last Glacial Maximum to the Holocene. *Nature*, **554**, 356 EP –, URL https://doi.org/10.1038/nature25454.

Rind, D., G. A. Schmidt, J. Jonas, R. Miller, L. Nazarenko, M. Kelley, and J. Romanski, 2018: Multicentury Instability of the Atlantic Meridional Circulation in Rapid Warming Simulations With GISS ModelE2. *Journal of Geophysical Research: Atmospheres*, **123 (12)**, 6331–6355, doi:10.1029/2017JD027149, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2017JD027149.

Rodgers, K. B., J. Lin, and T. L. Frölicher, 2015: Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosciences*, **12 (11)**, 3301–3320, doi:10.5194/bg-12-3301-2015, URL https://www.biogeosciences.net/12/3301/2015/.

Rohrschneider, T., B. Stevens, and T. Mauritsen, 2019: On simple representations of the climate response to external radiative forcing. *Climate Dynamics*, doi:10.1007/s00382-019-04686-4.

Rugenstein, M., and Coauthors, 2019: Equilibrium climate sensitivity estimated by equilibrating climate models. *in revision for GRL.*

Rugenstein, M. A. A., K. Caldeira, and R. Knutti, 2016a: Dependence of global radiative feedbacks on evolving patterns of surface heat fluxes. *Geophysical Research Letters*, **43 (18)**, 9877–9885, URL http://dx.doi.org/10.1002/2016GL070907.

Rugenstein, M. A. A., J. M. Gregory, N. Schaller, J. Sedláček, and R. Knutti, 2016b: Multiannual Ocean–Atmosphere Adjustments to Radiative Forcing. *Journal of Climate*, **29 (15)**, 5643–5659, URL http://dx.doi.org/10.1175/JCLI-D-16-0312.1.

Rugenstein, M. A. A., J. Sedláček, and R. Knutti, 2016c: Nonlinearities in patterns of long-term ocean warming. *Geophysical Research Letters*, **43 (7)**, 3380–3388, URL http://dx.doi.org/10.1002/2016GL068041.

Saint-Martin, D., and Coauthors, 2019: Fast forward to perturbed equilibrium climate. *Geophysical Research Letters*, **0 (ja)**, doi:10.1029/2019GL083031, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019GL083031, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019GL083031.

Salzmann, M., 2017: The polar amplification asymmetry: role of Antarctic surface height. *Earth System Dynamics*, **8 (2)**, 323–336, doi:10.5194/esd-8-323-2017, URL https://www.earth-syst-dynam.net/8/323/2017/.

Scheff, J., R. Seager, H. Liu, and S. Coats, 2017: Are Glacials Dry? Consequences for Paleoclimatology and for Greenhouse Warming. *Journal of Climate*, **30 (17)**, 6593–6609, doi:10.1175/JCLI-D-16-0854.1, URL https://doi.org/10.1175/JCLI-D-16-0854.1.

Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *Journal of Advances in Modeling Earth Systems*, **6 (1)**, 141–184, URL http://dx.doi.org/10.1002/2013MS000265.

Schneider, T., C. M. Kaul, and K. G. Pressel, 2019: Possible climate transitions from breakup of stratocumulus decks under greenhouse warming. *Nature Geoscience*, **12 (3)**, 163–167, doi:10.1038/s41561-019-0310-1, URL https://doi.org/10.1038/s41561-019-0310-1.

Senior, C. A., and J. F. B. Mitchell, 2000: The time-dependence of climate sensitivity. *Geophysical Research Letters*, **27 (17)**, 2685–2688, URL http://dx.doi.org/10.1029/2000GL011373.

Smith, R. S., J. M. Gregory, and A. Osprey, 2008: A description of the FAMOUS (version XD-BUA) climate model and control run. *Geoscientific Model Development*, **1 (1)**, 53–68, doi:10.5194/gmd-1-53-2008, URL https://www.geosci-model-dev.net/1/53/2008/.

Sniderman, J. M. K., and Coauthors, 2019: Southern Hemisphere subtropical drying as a transient response to warming. *Nature Climate Change*, doi:10.1038/s41558-019-0397-9, URL https://doi.org/10.1038/s41558-019-0397-9.

Song, X., and G. J. Zhang, 2014: Role of Climate Feedback in El Nino-like SST Response to Global Warming. *Journal of Climate*, doi:10.1175/JCLI-D-14-00072.1, URL http://dx.doi.org/10.1175/JCLI-D-14-00072.1.

Stouffer, R., and S. Manabe, 2003: Equilibrium response of thermohaline circulation to large changes in atmospheric CO_2 concentration. *Climate Dynamics*, **20 (7-8)**, 759–773, URL http://dx.doi.org/10.1007/s00382-002-0302-4.

Stouffer, R. J., and S. Manabe, 1999: Response of a Coupled Ocean–Atmosphere Model to Increasing Atmospheric Carbon Dioxide: Sensitivity to the Rate of Increase. *Journal of Climate*, **12 (8)**, 2224–2237, doi:10.1175/1520-0442(1999)012⟨2224:ROACOA⟩2.0.CO;2, URL http://dx.doi.org/10.1175/1520-0442(1999)012⟨2224:ROACOA⟩2.0.CO;2.

Svendsen, S. H., M. S. Madsen, Y. Suting, C. Rodehacke, and G. Adalgeirsdottir, 2015: An Introduction to the Coupled EC-Earth-PISM Model System. report 15-05, Danish Climate Centre.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2011: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93 (4)**, 485–498, doi:10.1175/BAMS-D-11-00094.1, URL http://dx.doi.org/10.1175/BAMS-D-11-00094.1.

Thomas, M. D., and A. V. Fedorov, 2019: Mechanisms and Impacts of a Partial AMOC Recovery Under Enhanced Freshwater Forcing. *Geophysical Research Letters*, **0 (0)**, doi:10.1029/2018GL080442, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL080442.

Trossman, D. S., J. B. Palter, T. M. Merlis, Y. Huang, and Y. Xia, 2016: Large-scale ocean circulation-cloud interactions reduce the pace of transient climate change. *Geophysical Research Letters*, **43 (8)**, 3935–3943, URL http://dx.doi.org/10.1002/2016GL067931.

Vial, J., J.-L. Dufresne, and S. Bony, 2013: On the interpretation of inter-model spread in CMIP5 climate sensitivity estimates. *Climate Dynamics*, **41 (11-12)**, 3339–3362, URL http://dx.doi.org/10.1007/s00382-013-1725-9.

Voldoire, A., and Coauthors, 2019: Evaluation of cmip6 deck experiments with cnrm-cm6-1. *Journal of Advances in Modeling Earth Systems*, **0 (0)**, doi:10.1029/2019MS001683, URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001683, https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019MS001683.

Voss, R., and U. Mikolajewicz, 2001: Long-term climate changes due to increased CO2 concentration in the coupled atmosphere-ocean general circulation model ECHAM3/LSG. *Climate Dynamics*, **17 (1)**, 45–60, doi:10.1007/PL00007925, URL https://doi.org/10.1007/PL00007925.

Winton, M., K. Takahashi, and I. M. Held, 2010: Importance of Ocean Heat Uptake Efficacy to Transient Climate Change. *Journal of Climate*, **23 (9)**, 2333–2344, doi:10.1175/2009JCLI3139.

1, URL http://dx.doi.org/10.1175/2009JCLI3139.1.

Yamamoto, A., A. Abe-Ouchi, M. Shigemitsu, A. Oka, K. Takahashi, R. Ohgaito, and Y. Yamanaka, 2015: Global deep ocean oxygenation by enhanced ventilation in the Southern Ocean under long-term global warming. *Global Biogeochemical Cycles*, **29 (10)**, 1801–1815, URL http://dx.doi.org/10.1002/2015GB005181.

Yeager, S. G., C. A. Shields, W. G. Large, and J. J. Hack, 2006: The Low-Resolution CCSM3. *Journal of Climate*, **19 (11)**, 2545–2566, URL http://dx.doi.org/10.1175/JCLI3744.1.

Yoshimori, M., M. Watanabe, H. Shiogama, A. Oka, A. Abe-Ouchi, R. Ohgaito, and Y. Kamae, 2016: A review of progress towards understanding the transient global mean surface temperature response to radiative perturbation. *Progress in Earth and Planetary Science*, **3 (1)**, 21, doi:10.1186/s40645-016-0096-3, URL https://doi.org/10.1186/s40645-016-0096-3.

Zhang, X., G. Lohmann, G. Knorr, and X. Xu, 2013: Different ocean states and transient characteristics in last glacial maximum simulations and implications for deglaciation. *Climate of the Past*, **9 (5)**, 2319–2333, doi:10.5194/cp-9-2319-2013, URL https://www.clim-past.net/9/2319/2013/.

Zhou, C., M. D. Zelinka, and S. A. Klein, 2016: Impact of decadal cloud variations on the Earth's energy budget. *Nature Geosci*, **9 (12)**, 871–874, URL http://dx.doi.org/10.1038/ngeo2828.

Zhu, J., Z. Liu, J. Zhang, and W. Liu, 2014: AMOC response to global warming: dependence on the background climate and response timescale. *Climate Dynamics*, **44 (11)**, 3449–3468, URL http://dx.doi.org/10.1007/s00382-014-2165-x.

Zickfeld, K., and Coauthors, 2013: Long-Term Climate Change Commitment and Reversibility: An EMIC Intercomparison. *Journal of Climate*, **26 (16)**, 5782–5809, URL http://dx.doi.org/10.1175/JCLI-D-12-00584.1.

LIST OF TABLES

TABLE 1. Description of collected variables. 2D means spatial resolution of latitude and longitude, except for *msftmyz* where it means latitude and depth. 3D means latitude, longitude, and depth. *msftmyz* is the sum of the eularian, eddybolus, and submeso component. For *so* and *thetao* there are also February and September values available for most models.

Shortname	Longname	Unit	Resolution
hfls	Surface Upward Latent Heat Flux	$W\,m^{-2}$	monthly, 2D
hfss	Surface Upward Sensible Heat Flux	$W\,m^{-2}$	monthly, 2D
pr	Precipitation on atmospheric grid	$kg\,m^{-2}\,s^{-1}$	monthly, 2D
psl	Sea Level Pressure	Pa	monthly, 2D
rlds	Surface Downwelling Longwave Radiation	$W\,m^{-2}$	monthly, 2D
rlus	Surface Upwelling Longwave Radiation	$W\,m^{-2}$	monthly, 2D
rlut	TOA Outgoing Longwave Radiation	$W\,m^{-2}$	monthly, 2D
rlutcs	TOA Outgoing Clear-Sky Longwave Radiation	$W\,m^{-2}$	monthly, 2D
rsds	Surface Downwelling Shortwave Radiation	$W\,m^{-2}$	monthly, 2D
rsdt	TOA Incident Shortwave Radiation	$W\,m^{-2}$	monthly, 2D
rsus	Surface Upwelling Shortwave Radiation	$W\,m^{-2}$	monthly, 2D
rsut	TOA Outgoing Shortwave Radiation	$W\,m^{-2}$	monthly, 2D
rsutcs	TOA Outgoing Clear-Sky Shortwave Radiation	$W\,m^{-2}$	monthly, 2D
tas	Near-Surface Air Temperature	K	monthly, 2D
ts	Atmospheric surface temperature	K	monthly, 2D
sic	Sea Ice Area Fraction	%	monthly, 2D
msftmyz	Meridional Overturning Circulation	$m^3\,s^{-1}$	annual, 2D
tos	Sea surface temperature	K	annual, 2D
sos	Sea surface salinity	psu	annual, 2D
wfo	Net water flux into sea water	$kg\,m^{-2}\,s^{-1}$	annual, 2D
evs	Water evaporation	$kg\,m^{-2}\,s^{-1}$	annual, 2D
pr_ocn	Precipitation (rain and snow) on ocean grid	$kg\,m^{-2}\,s^{-1}$	annual, 2D
tauuo	Surface downward wind stress in x direction	$N\,m^{-2}$	annual, 2D
tauvo	Surface downward wind stress in y direction	$N\,m^{-2}$	annual, 2D
so	Sea Water Salinity	psu	annual, 3D
thetao	Sea Water Potential Temperature	K	annual, 3D

TABLE 2. Overview of models and contributed simulations. The resolution of atmosphere and ocean is given in # of grid points per latitude x longitude, and latitude x longitude x depth, respectively. Models are referred to by their shortnames throughout the manuscript. Section 2b describes the forcing levels. References in the last column describe the models and simulations. Some simulations are published in their full length, some simulations contributed to LongRunMIP are the extensions of simulations discussed in the references, and some simulations are unpublished.

Model (shortname)	**Forcing level shortname**	**Length (yrs)**	**Atmosphere resolution**	**Ocean resolution**	**Control sim (yrs)**	**Model and simulation documentation**
CCSM3 **CCSM3**	abrupt2x abrupt4x abrupt8x	3000 2120 1450	48 x 96	100 x 116 x 25	1530	Yeager et al. (2006) Danabasoglu and Gent (2009)
CCSM3 **CCSM3II**	abrupt2.4 abrupt4.8 lin2.4	3701 3132 3990	48 x 96	100 x 116 x 25	3805	Yeager et al. (2006) Castruccio et al. (2014)
CESM 1.0.4 **CESM104**	abrupt2x abrupt4x abrupt8x	2500 5900 5100	96 x 144	384 x 20 x 60	1320	Gent et al. (2011) Danabasoglu et al. (2012) Rugenstein et al. (2016c)
CNRM-CM6-1 **CNRMCM61**	abrupt2x abrupt4x	750 1850	128 x 256	180 x 360 x 75	2000	Voldoire et al. (2019) Saint-Martin et al. (2019)
EC-Earth-PISM **ECEARTH**	historical RCP8.5+	1270	160 x 320	292 x 362 x 42	508	Hazeleger et al. (2012) Svendsen et al. (2015)
ECHAM5/MPIOM **ECHAM5**	abrupt4x 1pct4x	1000 6080	48 x 96	101 x 120 x 40	100	Jungclaus et al. (2006) Li et al. (2013)
FAMOUS **FAMOUS**	abrupt2x abrupt4x	3000 3000	37 x 48	73 x 96 x 20	3000	Smith et al. (2008)
GFDL-CM3 **GFDLCM3**	1pct2x	5000	90 x 144	200 x 360 x 50	5200	Donner et al. (2011) Paynter et al. (2018)
GFDL-ESM2M **GFDLESM2M**	1pct2x	4500	90 x 144	200 x 360 x 50	1340	Dunne et al. (2012) Paynter et al. (2018)
GISS-E2-R **GISSE2R**	abrupt4x 1pct4x	5000 5000	90 x 144	180 x 288 x 32	5225	Schmidt et al. (2014); Miller et al. (2014); Nazarenko et al. (2015) Rind et al. (2018)
HadCM3L **HadCM3L**	abrupt2x abrupt4x abrupt6x abrupt8x	1000 1000 1000 1000	73 x 96	73 x 96 x 20	1000	Cox et al. (2000) Cao et al. (2016)
HadGEM2-ES **HadGEM2**	abrupt4x	1328	145 x 192	216 x 360 x 40	239	Collins et al. (2011) Andrews et al. (2015)
IPSL-CM5A-LR **IPSLCM5ALR**	abrupt4x	1000	96 x 96	149 x 182 x 31	1000	Dufresne et al. (2013)
MIROC 3.2 **MIROC32**	1pct2x 1pct4x	2000 2000	64 x 128	192 x 256 x 44	681	Hasumi and Emori (2004) Yamamoto et al. (2015); Yoshimori et al. (2016)
MPIESM-1.2 **MPIESM12**	abrupt2x abrupt4x abrupt8x abrupt16x	1000 1000 1000 1000	96 x 192	220 x 256 x 40	1237	Mauritsen et al. (2018) Rohrschneider et al. (2019)
MPIESM-1.1 **MPIESM11**	abrupt4x	4459	96 x 192	220 x 256 x 40	2000	Mauritsen et al. (2018)

TABLE 3. Published millennial-length simulations

Paper	Model	Forcing level	Length (yr)	Content/scientific comment
Senior and Mitchell (2000)	HadCM2	2xCO_2	$\approx$ 800	Included flux adjustments; effective climate sensitivity increases due to SW CRE due to changes in the inter-hemispheric temperature gradient
Bi et al. (2001)	CSIRO	3xCO_2	$\approx$ 1000	Cessation and recovery of Antarctic Bottom Water and North Atlantic Deep Water formation
Voss and Mikolajewicz (2001)	ECHAM3	2x and 4xCO_2	850	Adjustment time scales, committed warming, ocean thermohaline circulation
Stouffer and Manabe (1999, 2003)	GFDL	0.5x, 2x, and 4xCO_2	4000	Thermohaline circulation and paleo-oceanographic implications
Boer and Yu (2003b,a,c)	CCCma	2050 and 2100 forcing	1000	Radiative feedbacks and surface warming; effective climate sensitivity decreases with time; slab versus fully coupled models
Gregory et al. (2004)	HadCM3	2xCO_2	$\approx$ 1000	TOA radiative imbalance and surface temperature are not linearly related; after 1000 yr the model is still 0.7 W m^{-2} away from equilibrium
* Danabasoglu and Gent (2009)	CCSM3	2x, 4x, and 8xCO_2	3000	Comparing slab and fully coupled models; determining ECS; Jonko et al. (2013) analyzed the contributions of different feedbacks to doublings of CO_2
Gillett et al. (2011)	CanESM1	21st century	$\approx$ 1000	Impact of reduced emissions
* Li et al. (2013)	ECHAM5/MPI-OM	2xCO_2	$\approx$ 6000	Comparing slab and fully coupled models; determining ECS; adjustment time scales of surface warming patterns, ocean heat uptake, and sea level rise
Frölicher et al. (2014); Frölicher and Paynter (2015)	GFDL-ESM2M, CSM1	4xCO_2 pulse	1000	Climate impact of CO_2 emission stoppage; evolving feedbacks; ECS; transient climate response to cumulative carbon emissions
* Andrews et al. (2015)	HadGEM2-ES	4xCO_2	$\approx$ 1300	Non-constancy of feedbacks; variations of TOA components cancel each other on the century to millennial time scale
* Yamamoto et al. (2015); Yoshimori et al. (2016)	MIROC 3.2	2x and 4xCO_2	2000	Deep ocean ventilation overall increases oxygenation after a transient decrease; review article on ocean heat uptake in coupled models and energy balance models
* Cao et al. (2016)	HadCM3L	2x, 4x, 6x, 8xCO_2	1000	Comparing CO_2 to other forcing agents and geo-engineering scenarios
* Rugenstein et al. (2016b,a)	CESM104	2x, 4x, 8xCO_2	$\approx$ 1300	Dependence of global and regional radiative feedback evolution on surface heat flux patters; forcing adjustment
* Paynter et al. (2018)	GFDL-ESM2M, GFDL-CM3	2xCO_2	$\approx$ 5000	Evolution of global and regional radiative feedbacks and the role of atmospheric vertical velocity fields and inversion strengths
* Rind et al. (2018)	GISS-E2-R	4xCO_2	$\approx$ 2000	AMOC reduction and recovery on North Atlantic surface flux conditions
Krasting et al. (2018)	GFDL-ESM2Mb, GFDL-ESM2G	4xCO_2	5000	Ocean heat uptake, model formulation of diapycnal diffusivity and ocean vertical coordinates

LIST OF FIGURES

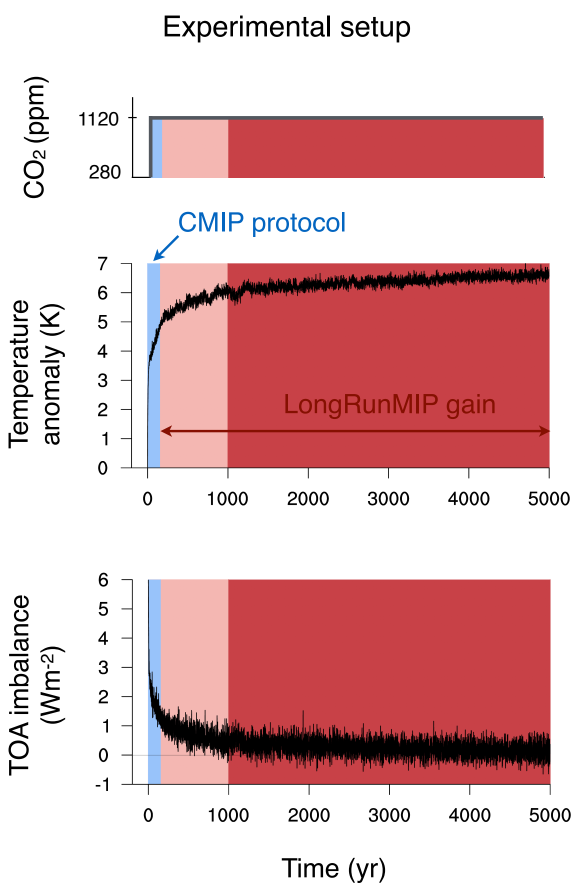


FIG. 1. Global and annual mean surface air temperature (*tas* in Table 1) anomaly and top of the atmosphere (TOA) radiative imbalance (computed as *rsdt* - *rlut* - *rsut*, see Table 1) to a step-forcing of quadrupling CO_2 as simulated by the CESM104 model. For the Coupled Model Intercomparison Project Phase 5 and 6, this simulation is part of the standard protocol, but only 150 simulated years are requested (blue shading). We collect simulations that extended this experiment for at least 850 years (light red shading), ideally until they are equilibrated (end of dark red shading).

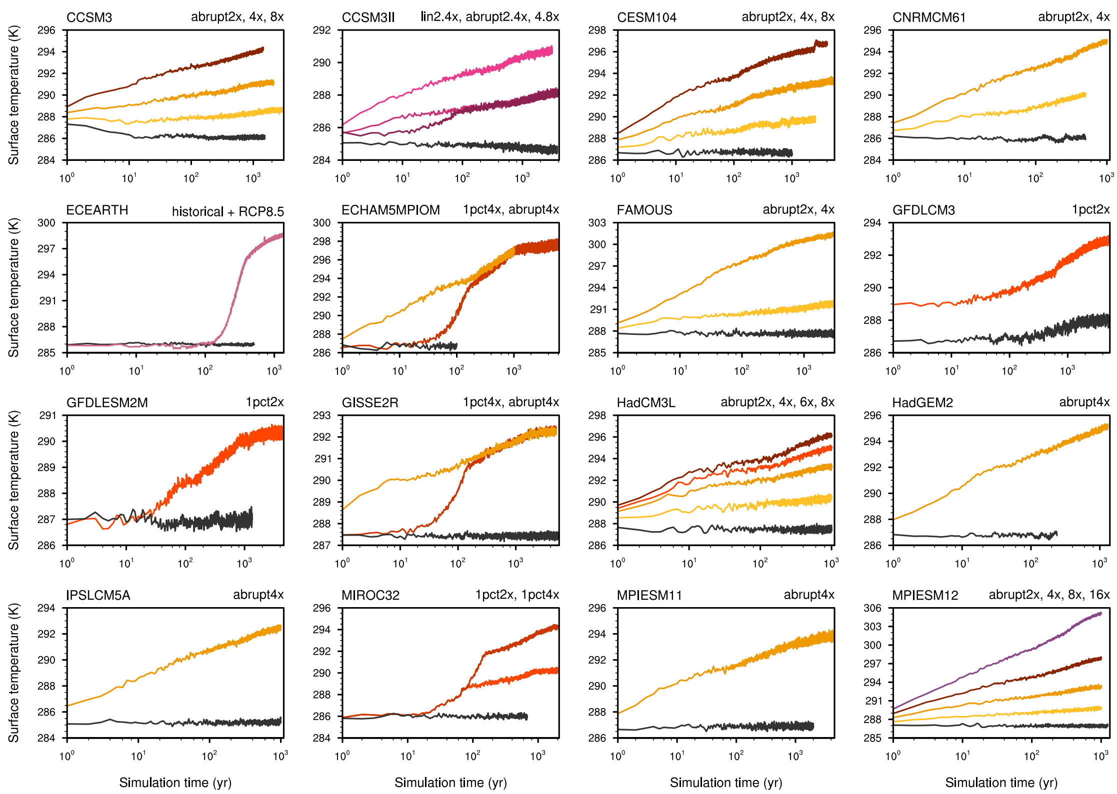


FIG. 2. Global annual mean surface air temperature for all control (black) and forced (color, listed in the top right of each panel) simulations. *abrupt2x*, *4x*, *6x*, *8x* means that the CO_2 concentration is doubled, quadrupled, sextupled, octupcliated, as a step-forcing branched off the control simulation. *1pct2x* and *1pct4x* means the CO_2 concentration is linearly increased 1 % per year until the concentration is doubled or quadrupled, respectively. The simulations of ECEARTH and CCSM3II are described in Section b. Note the different axis ranges for each model. GFDLCM3 and CCSM3II are not branched off directly from the control simulation.

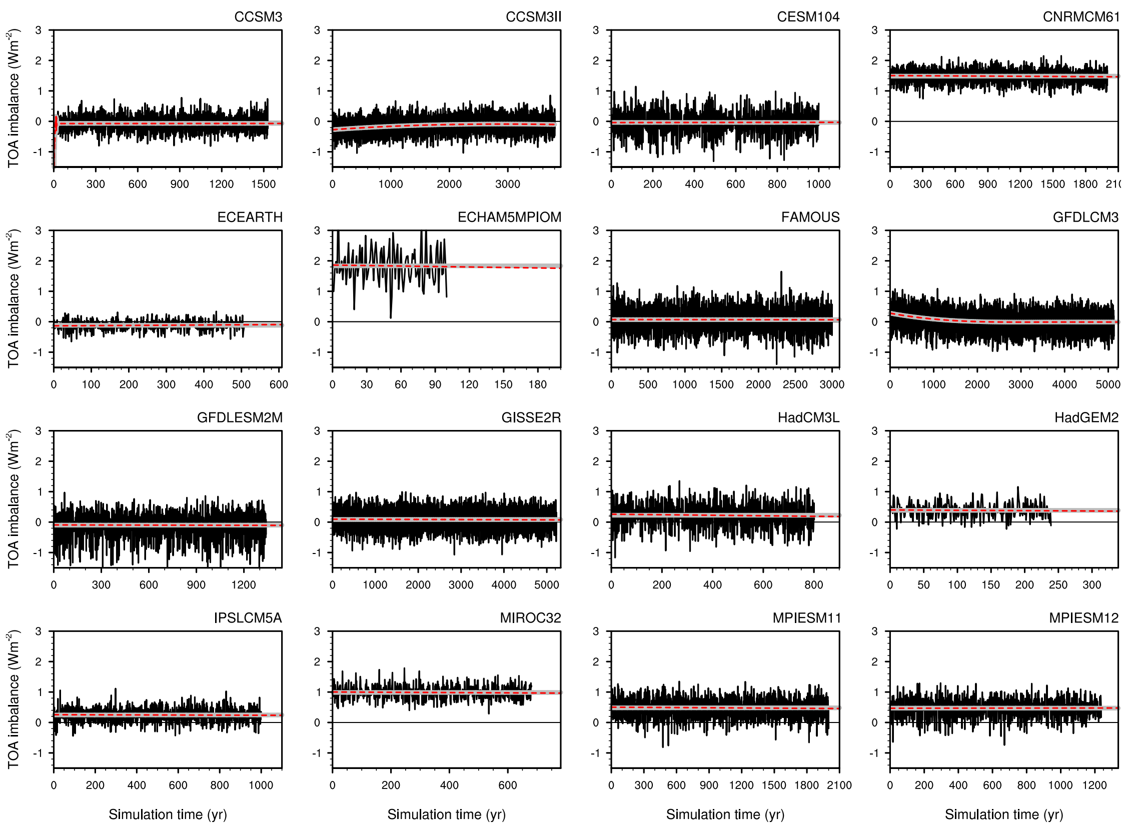
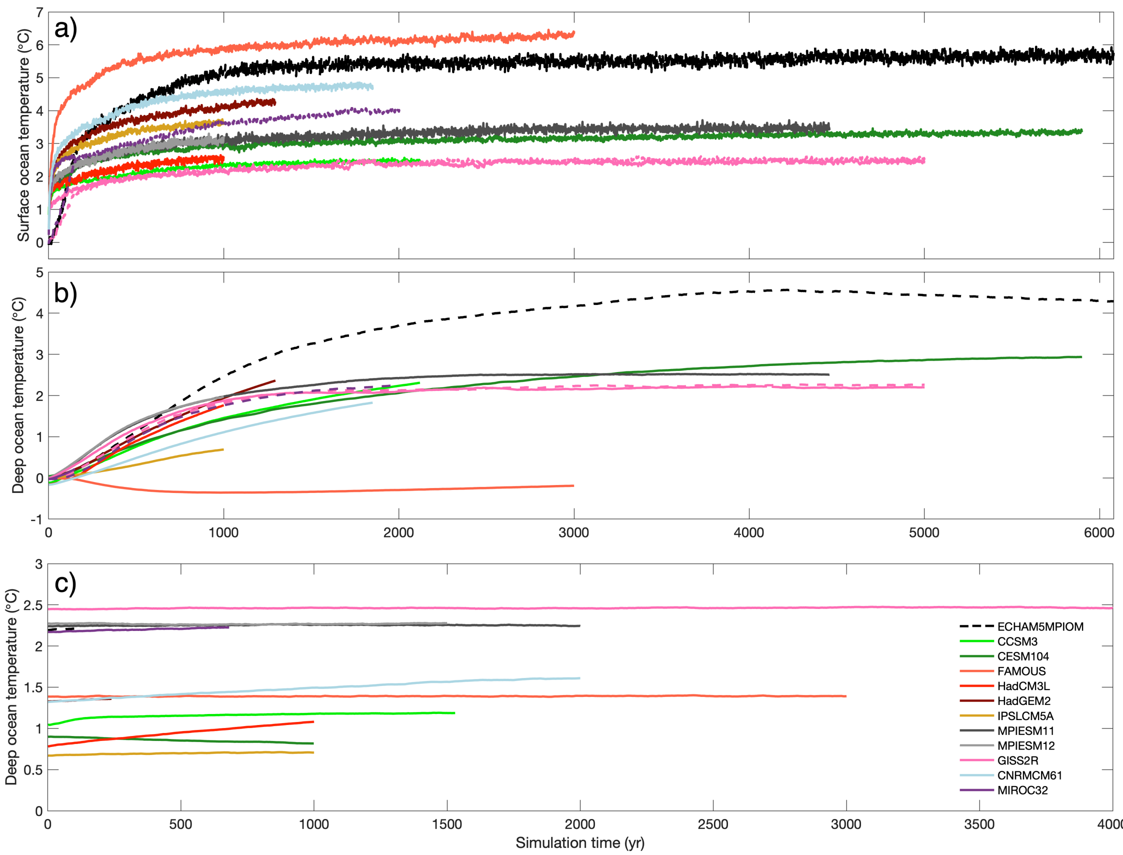


FIG. 3. Top of the atmosphere (TOA) annual and global mean radiative imbalance of all control simulations. Note the different lengths of the horizontal axes. The gray line indicates the average, the red line the linear trend, except for CCSM3II and GFDLCM3 for which both colors depict a fourth-order-polynomial fit.



890 FIG. 4. Global and annual mean temperature anomalies (experiment minus average of the control simulation) of the surface ocean (a, first layer) and deep ocean (b), as well as absolute values of deep ocean temperature in the control simulations (c), for *abrup4x* (solid) and *1pct4x* (dashed) simulations. "Deep ocean" means around 2000 m depth (closest level). Note that the time scale in c) is shorter than in a) and b).

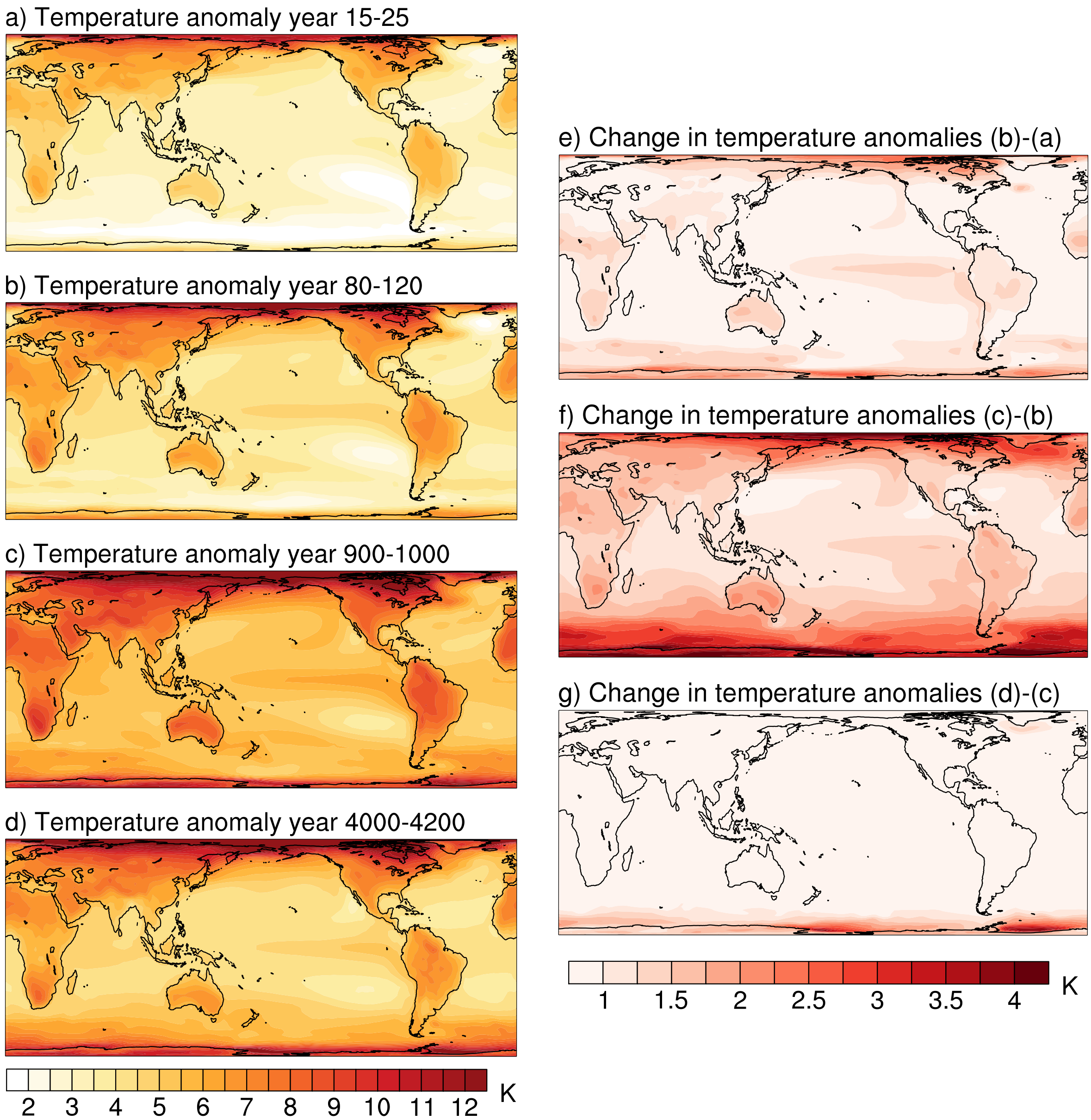


FIG. 5. Time evolution of the surface air temperature anomaly in the *abrupt4x* simulations. The model mean of CCSM3, CESM104, CNRMCM61, ECHAM5, GISSE2R, HadCM3L, HadGEM2, IPSLCM5A, MPIESM11, and MPIESM12 is shown in panel a, b, c, e, and f, while the model mean of only CESM104, GISSE2R, and MPIESM11 is shown in panel d and g, due to the length of these contributions. See Table 2 for details of the length of each simulation.

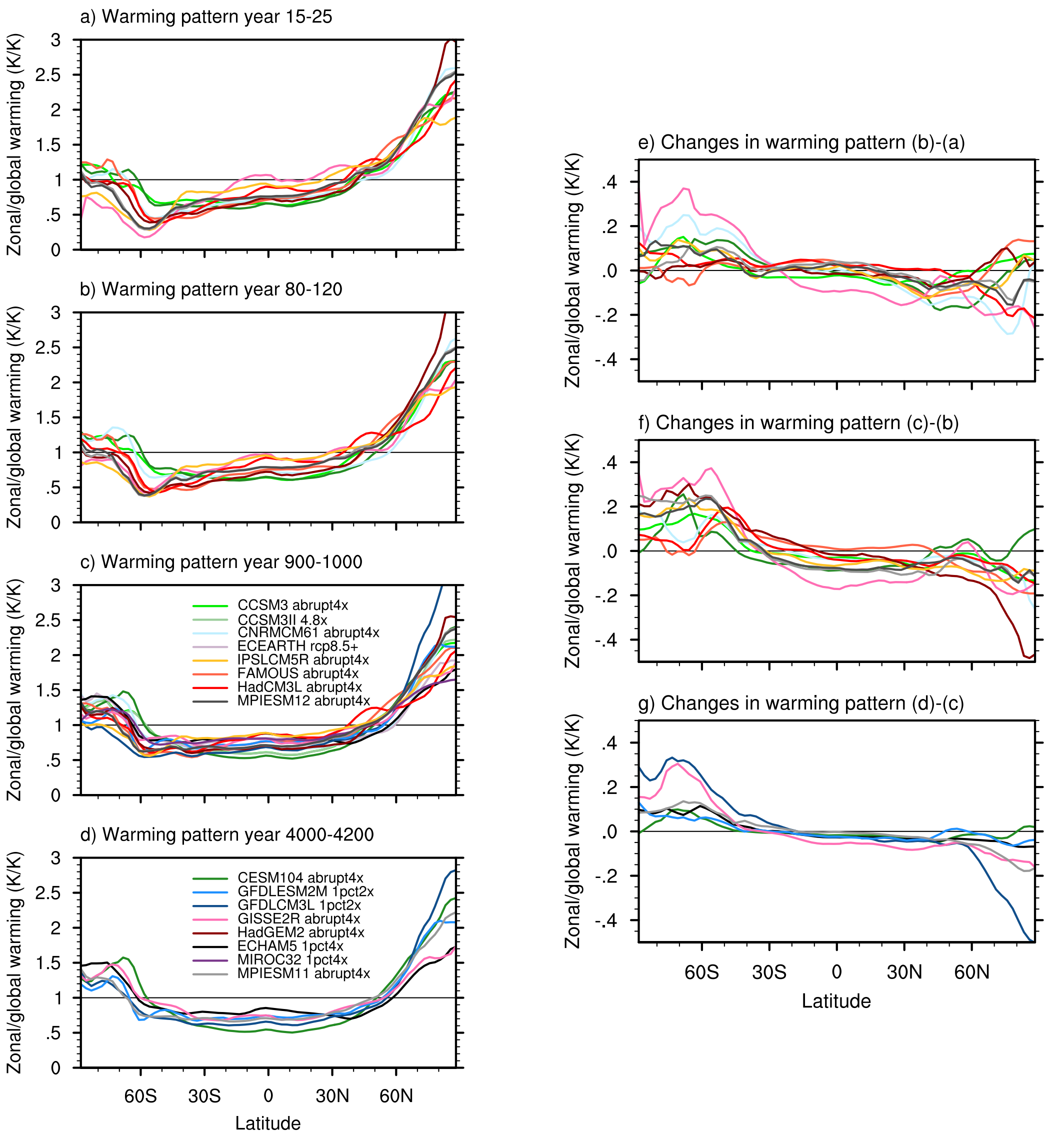


FIG. 6. Time evolution of the zonal mean surface air temperature response normalized by the global mean temperature anomaly. Above (below) 1 means that warming is amplified (reduced) relative to the globally mean warming (a-d). Panel e-g show the differences (note the difference scale). Panel a, b, e, and f contain only *abrupt4x* simulations, while panel c, d, and g also contain the *1pct2x* and *RCP8.5+* simulations with integration lengths above 4000 years. Table 2 lists all simulations and model long names.

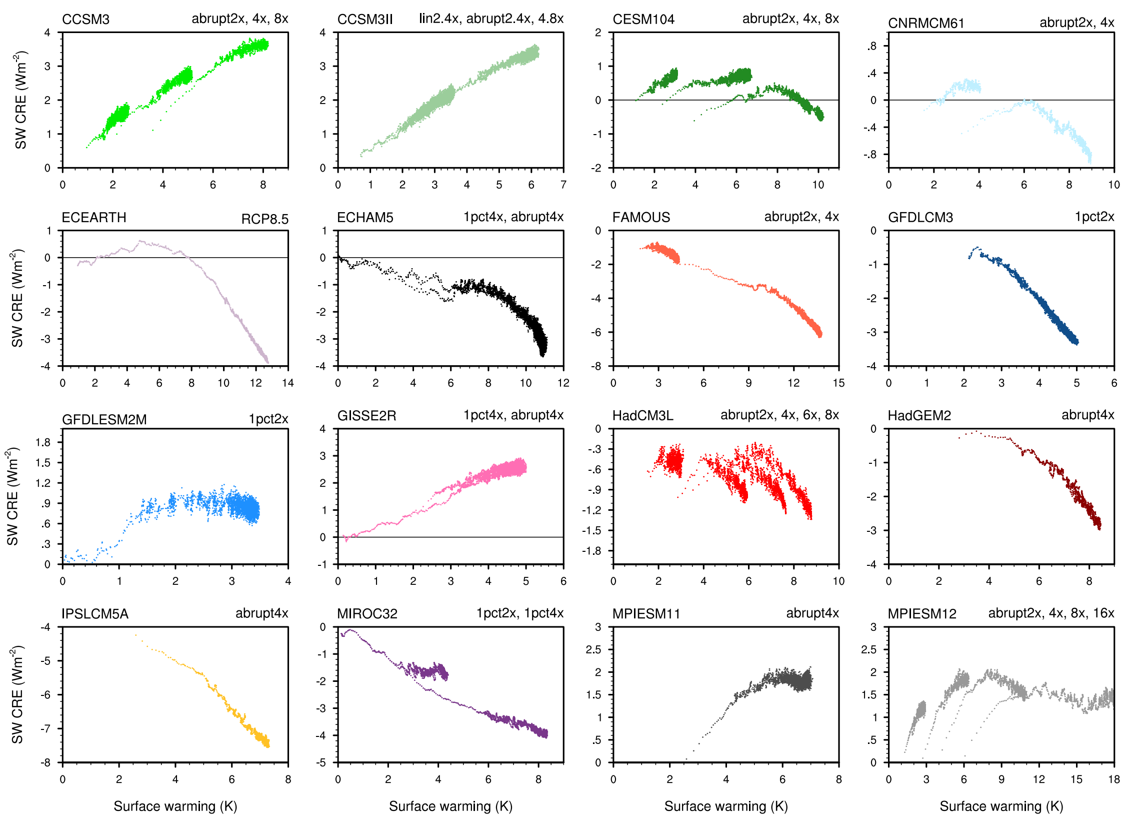


FIG. 7. Simulated shortwave cloud radiative effects SW CRE for different levels of global surface air temperature changes. Each point is a ten-year running average. Note the different axes labels, which cover a large range in TOA imbalance and surface temperature. Table 2 lists all simulations and model long names.